



Earth Observing System Clearinghouse

ECHO 10.0

Metadata Requirements and Recommendations

Document Version 10.5

REVISION HISTORY

Date	Version	Brief Description	Author
12/11/07	10.1	Initial version for ECHO 10	Matt Cechini
12/11/07	10.2	Reformat to be consistent with other ECHO documentation. Minor editorial changes.	Judith Wright
12/12/07	10.3	Updated page breaks between sections	Matt Cechini
12/14/07	10.4	Added the AssociatedDIFs to the list of recommended collection metadata elements.	Matt Cechini
12/14/07	10.5	Changed to Requirements and Recommendations	Matt Cechini

TABLE OF CONTENTS

1. OVERVIEW.....	1
2. COLLECTION METADATA	1
2.1. Required Fields	1
2.2. Recommended Fields.....	2
2.3. Sample Metadata File	4
3. GRANULE METADATA.....	7
3.1. Required Fields	7
3.2. Recommended Fields.....	7
3.3. Sample Metadata File	9
4. BROWSE METADATA.....	11
4.1. Required Fields	11
4.2. Recommended Fields.....	11
4.3. Sample Metadata File	11

LIST OF TABLES

TABLE 1. COLLECTION REQUIRED FIELDS	1
TABLE 2. COLLECTION RECOMMENDED FIELDS	2
TABLE 3. GRANULE REQUIRED FIELDS	7
TABLE 4. GRANULE RECOMMENDED FIELDS	7
TABLE 5. BROWSE REQUIRED FIELDS.....	11
TABLE 6. BROWSE RECOMMENDED FIELDS.....	11

ECHO 10.0 Metadata Requirements and Recommendations

1. OVERVIEW

The purpose of this document is to give an overview of the metadata requirements and recommendations for ingesting science metadata into the ECHO 10.0.x system. The following two perspectives are presented during this analysis:

- **Required Fields** – The metadata elements that must be present in order to pass the base XML Schema validation
- **Recommended Fields** – The metadata elements that we recommend to facilitate searching and data usability by the science community

The metadata name and description of each metadata field are listed in tabular form, along with the data type and relevant constraints, where applicable. Only the top-level elements of complex metadata elements are listed. Where this occurs, the requisite child element information is inherently required to correctly represent the parent metadata element. Accompanying each metadata type's requirements and recommendations is a sample Ingest metadata file that implements the outlined requirements.

2. COLLECTION METADATA

2.1. Required Fields

Table 1. Collection Required Fields

Name	Data Type	Constraint	Description
ShortName	String	MaxLength = 85	This attribute identifies the short name associated with the collection. This is the official reference name used in identifying the contents of the data collection. All characters must be in upper case.
VersionId	String	MaxLength = 80	This attribute specifies the version identifier of the data collection
InsertTime	DateTime		This attribute specifies the insert date/time the collection entered the data provider's database. This date is provided by the data provider.
LastUpdate	DateTime		This attribute specifies the most recent date/time an update occurred in the data provider's database. This date is provided by the data provider.
LongName	String	MaxLength = 1024	This attribute will identify the long name associated with the collection. This is the reference name used in describing the scientific contents of the data collection.
DataSetId	String	MaxLength = 1030	This attribute specifies a unique name for the collection. This information is computed by ECHO according to the data provider's policy.
Description	String	MaxLength = 4000	This attribute identifies the major emphasis of the content of the collection.

Name	Data Type	Constraint	Description
Orderable	Boolean		This attribute indicates whether this collection is orderable.
Visible	Boolean		<p>This attribute indicates whether the collection is visible. Visibility is a basic access control mechanism that bypasses all ACL rules. If a collection is not visible, only users with the owning provider role will be able to see the item. No other users will be able to see the item—no matter what ACL permissions are in place.</p> <p>If group based permissions are needed, use the Restriction Flag field instead of visibility.</p> <p>Normally, visibility is used when an item is first ingested. Making the item <i>not visible</i> allows the provider time to install any ACL rules or order options required before users will see the item.</p> <p>Visibility is more commonly set at the collection level than the granule level. If a collection is not visible, none of the granules in the collection will be visible.</p>

2.2. Recommended Fields

Table 2. Collection Recommended Fields

Name	Data Type	Constraint	Description
ProcessingLevelId	String	MaxLength = 80	The processing level class contains the level identifier and level description of the collection.
Price	Decimal		This attribute specifies the price for ordering the collection
SpatialKeywords			This attribute specifies a word or phrase that serves to summarize the spatial regions covered by the collection. It may be repeated if several regions are covered. This often occurs when a collection is described as covering some large region, and several smaller sub regions within that region.
TemporalKeywords			This attribute specifies a word or phrase that serves to summarize the temporal characteristics referenced in the collection.
Temporal			This attribute contains records that describe the basis of the time system used for a specific collection.
Contact			This attribute contains the basic characteristics for a person or an organization type of contact. These contacts may provide information about a Collection, Delivered Algorithm Package, PGE, or Data Originator. System and user profile contact information is held elsewhere.
DisciplineTopicParameters			This attribute holds a cross reference between collections and discipline keywords. The cross reference could further involve associations between discipline keywords and a collection's platform/instrument configuration.

Name	Data Type	Constraint	Description
Platform			This attribute describes the relevant platforms associated with the acquisition of the collection or granule. Platform types include Spacecraft, Aircraft, Vessel, Buoy, Platform, Station, Network, Human etc.
Instrument			This entity registers the device used to measure or record data, including direct human observation. In cases where instruments have a single sensor or the instrument and sensor are used synonymously (e.g., AVHRR), both the Instrument and Sensor should be recorded. The Sensor information is represented by other entities.
Sensor			This entity holds the referential information for the collection source/sensor configuration including sensor parameter settings such as technique, etc.
Campaigns			This entity contains attributes describing the scientific endeavor(s) to which the collection is associated. Scientific endeavors include campaigns, projects, interdisciplinary science investigations, missions, field experiments, etc.
OnlineAccessURL			This entity stores the online URL(s) for the granule, if there any. The URL either provides the site the user can obtain granule data or gives further instructions for obtaining the granule data.
OnlineResource			This entity records the documentation information of the collection including documentation type and documentation URL where applicable. These resources may include additional subsetting or processing services available for the granule.
AssociatedDIFs		MaxLength = 80 MinLength = 1	This entity specifies the collection's Directory Interchange Format (DIF) identifier. The DIF identifier is used to uniquely identify a provider's data set. ECHO coordinates this list with the GCMD portal to facilitate the discovery of collections through GCMD. Providers may specify their own format, but the following is recommended: <ShortName>_V<Version_ID> (ex: MOD14_V005). Notice that the version id is zero padded to be three digits long.
Spatial			This entity contains the collection's spatial coverage information.

2.3. Sample Metadata File

```
<?xml version="1.0" encoding="utf-8"?>
<CollectionMetaDataTable>

<Collections>
    <Collection>
        <ShortName>MISR_ABC</ShortName>
        <VersionId>1</VersionId>
        <InsertTime>2007-11-19T10:00:00Z</InsertTime>
        <LastUpdate>2007-11-19T12:00:00Z</LastUpdate>
        <LongName>MISR_ABC LongName</LongName>
        <DataSetId>MISR ABC DataSetID</DataSetId>
        <Description>This is a sample dataset.</Description>
        <Orderable>true</Orderable>
        <Visible>false</Visible>
        <ProcessingLevelId>1</ProcessingLevelId>
        <Price>100.00</Price>

        <SpatialKeywords>
            <Keyword>SpatialKeyword1</Keyword>
        </SpatialKeywords>

        <TemporalKeywords>
            <Keyword>TemporalKeyword1</Keyword>
        </TemporalKeywords>

        <Temporal>
            <RangeDateTime>
                <BeginningDateTime>2007-04-01T01:00:00Z</BeginningDateTime>
                <EndingDateTime>2007-05-01T01:00:00Z</EndingDateTime>
            </RangeDateTime>
        </Temporal>

        <Contacts>
            <Contact>
                <Role>User Services</Role>
                <OrganizationPhoneNumbers>
                    <PhoneNumber>
                        <PhoneNumber>555-555-5555</PhoneNumber>
                        <Type>Primary</Type>
                    </PhoneNumber>
                </OrganizationPhoneNumbers>
                <OrganizationEmails>
                    <Email>email@domain.com</Email>
                </OrganizationEmails>
                <ContactPersons>
                    <ContactPerson>
                        <FirstName>Bob</FirstName>
                        <LastName>Orbit</LastName>
                    </ContactPerson>
                </ContactPersons>
            </Contact>
        </Contacts>
    </Collection>
</Collections>
</CollectionMetaDataTable>
```

```
<DisciplineTopicParameters>
  <DisciplineTopicParameter>
    <DisciplineKeyword>Earth Science</DisciplineKeyword>
    <TopicKeyword>Atmosphere</TopicKeyword>
    <TermKeyword>Atmospheric Pressure</TermKeyword>
    <VariableKeyword>Atmospheric Pressure Measurements</VariableKeyword>
  </DisciplineTopicParameter>
</DisciplineTopicParameters>

<Platforms>
  <Platform>
    <ShortName>Aura</ShortName>
    <LongName>EOS Aura Mission Satellite</LongName>
    <Type>Spacecraft</Type>
    <Characteristics>
      <Characteristic>
        <Name>OrbitInclination</Name>
        <Description>
          Angle between the orbit plane and the Earth's
          equatorial plane
        </Description>
        <DataType>float</DataType>
        <Unit>Degrees</Unit>
        <Value>98.2</Value>
      </Characteristic>
    </Characteristics>
    <Instruments>
      <Instrument>
        <ShortName>HIRDLS</ShortName>
        <LongName>High Resolution Dynamic Limb Sounder</LongName>
        <Sensors>
          <Sensor>
            <ShortName>HRDLS Channel 21</ShortName>
            <LongName>
              HRDLS Channel 21 NO2 6.12-6.32
            </LongName>
            <Technique>Infrared Limb Sounding</Technique>
          </Sensor>
        </Sensors>
      </Instrument>
    </Instruments>
  </Platform>
</Platforms>

<Campaigns>
  <Campaign>
    <ShortName>HRDLS</ShortName>
    <LongName>
      High Resolution Dynamics Limb Sounder (EOS CHEM Experiment)
    </LongName>
  </Campaign>
</Campaigns>
```

```
<OnlineAccessURLs>
  <OnlineAccessURL>
    <URL>Text</URL>
    <URLDescription>Text</URLDescription>
    <MimeType>Text</MimeType>
  </OnlineAccessURL>
</OnlineAccessURLs>

<OnlineResources>
  <OnlineResource>
    <URL>http://aura.gsfc.nasa.gov/instruments/hirdls/</URL>
    <Description>
      This URL contains information and data access for HRDLS
      and all the Goddard DAAC's Aura Instruments.
    </Description>
    <Type>ASSOCIATED DATA SET(S)</Type>
  </OnlineResource>
</OnlineResources>

<AssociatedDIFs>
  <EntryId>MOD14_V005</EntryId>
</AssociatedDIFs>

<Spatial>
  <SpatialCoverageType>BOTH</SpatialCoverageType>
  <HorizontalSpatialDomain>
    <!--Global>??</Global-->
    <ZoneIdentifier></ZoneIdentifier>
    <Geometry>
      <CoordinateSystem>CARTESIAN</CoordinateSystem>
      <BoundingRectangle>
        <WestBoundingCoordinate>-180.0</WestBoundingCoordinate>
        <NorthBoundingCoordinate>38.0</NorthBoundingCoordinate>
        <EastBoundingCoordinate>180.0</EastBoundingCoordinate>
        <SouthBoundingCoordinate>-38.0</SouthBoundingCoordinate>
      </BoundingRectangle>
    </Geometry>
  </HorizontalSpatialDomain>
  <VerticalSpatialDomain>
    <Type>Atmosphere Layer</Type>
    <Value>Atmosphere Profile</Value>
  </VerticalSpatialDomain>
  <GranuleSpatialRepresentation>CARTESIAN</GranuleSpatialRepresentation>
</Spatial>

  </Collection>
</Collections>
</CollectionMetaDataFile>
```

3. GRANULE METADATA

3.1. Required Fields

Table 3. Granule Required Fields

Name	Data Type	Constraint	Description
GranuleUR	String	MaxLength = 250 MinLength = 1	The Universal Reference (UR) ID of the granule referred to by the data provider. This ID is unique per data provider.
InsertTime	DateTime	None	This attribute is the date/time this granule entered the data provider's database.
LastUpdate	DateTime	None	This attribute is the date/time the data provider last updated the granule information in the data provider's database
Collection	N/A	N/A	This attribute holds the reference information for a granule to a collection. The Granule references the collection either by collection short name and collection version or by collection dataset ID.
Orderable	Boolean	None	This attribute indicates whether the granule is orderable.

3.2. Recommended Fields

Table 4. Granule Recommended Fields

Name	Data Type	Constraint	Description
DataGranule			This entity stores the basic descriptive characteristics associated with a granule.
PGEVersionClass			This entity stores basic descriptive characteristics related to the Program Generation Executable (PGE) associated with a granule.
Temporal			This entity contains records that describe the basis of the time system used for a specific collection.
Spatial			This entity contains the granule's spatial coverage information.
MeasuredParameters			<p>This attribute specifies the names of the geophysical parameters expressed in the data as well as associated quality flags and quality status.</p> <p>The quality status contains measures of quality for the granule. The parameters used to set these measures are not preset and will be determined by the data producer. Each set of measures can occur many times either for the granule as a whole or for individual parameters.</p> <p>The quality flags contain the science, operational, and automatic quality flags that indicate the overall quality assurance levels of specific parameter values within a granule.</p>
Platform			This attribute describes the relevant platforms associated with the acquisition of the collection or granule. Platform types include Spacecraft, Aircraft, Vessel, Buoy, Platform, Station, Network, Human, etc.

Name	Data Type	Constraint	Description
Instrument			This attribute registers the device used to measure or record data, including direct human observation. In cases where instruments have a single sensor or the instrument and sensor are used synonymously (e.g., AVHRR) both the Instrument and Sensor should be recorded. The Sensor information is represented by other entities.
Sensor			This attribute holds the referential information for the granule source/sensor configuration including sensor parameter settings such as technique, etc.
Campaigns			This entity contains attributes describing the scientific endeavor(s) to which the granules is associated. Scientific endeavors include campaigns, projects, interdisciplinary science investigations, missions, field experiments, etc.
DataFormat	String	MaxLength=80	This attribute contains the file format of the raw data (such as HDF) for this granule
Price	Decimal		This attribute specifies the price of the granule data when ordered.
OnlineAccessURL			This entity stores the online URL(s) for the granule, if any. The URL either provides the site the user can obtain granule data or gives further instructions for obtaining the granule data.
OnlineResource			This entity records the documentation information of the granule including documentation type and documentation URL where applicable. These resources may include additional subsetting or processing services available for the granule.
CloudCover	Decimal		A percentage value indicating how much of the area of a granule (the ECS data unit) has been obscured by clouds. It is worth noting that there are many different measures of cloud cover within the ECS data holdings and that the cloud cover parameter that is represented in the DataPool is specific to Data Set.
AssociatedBrowselImages			The list of associated browse images to this granule

3.3. Sample Metadata File

```
<?xml version="1.0" encoding="utf-8"?>
<GranuleMetaDataTable>

    <Granules>
        <Granule>
            <GranuleUR>AIRI2CCF.005:AIRS.2004.07. G07270040936.hdf</GranuleUR>
            <InsertTime>2007-09-27T08:10:10Z</InsertTime>
            <LastUpdate>2007-09-30T08:10:10Z</LastUpdate>

            <Collection>
                <ShortName>MISR_ABC</ShortName>
                <VersionId>1</VersionId>
                <DataSetId>MISR ABC DataSetID</DataSetId>
            </Collection>

            <DataGranule>
                <SizeMBDataGranule>8713694</SizeMBDataGranule>
                <DayNightFlag>NIGHT</DayNightFlag>
                <ProductionDateTime>2007-10-02T09:00:00Z</ProductionDateTime>
            </DataGranule>

            <PGEVersionClass>
                <PGEName>Name</PGEName>
                <PGEVersion>05</PGEVersion>
            </PGEVersionClass>

            <Temporal>
                <SingleDateTime>2007-04-21T01:00:00Z</SingleDateTime>
            </Temporal>

            <Spatial>
                <VerticalSpatialDomains>
                    <VerticalSpatialDomain>
                        <Type>Maximum Altitude</Type>
                        <Value>TOA</Value>
                    </VerticalSpatialDomain>
                </VerticalSpatialDomains>
                <HorizontalSpatialDomain>
                    <Geometry>
                        <BoundingRectangle>
                            <WestBoundingCoordinate>
                                75.2058334350586
                            </WestBoundingCoordinate>
                            <NorthBoundingCoordinate>
                                -31.4805202484131
                            </NorthBoundingCoordinate>
                            <EastBoundingCoordinate>
                                104.000869750977
                            </EastBoundingCoordinate>
                            <SouthBoundingCoordinate>
                                -55.2935943603516
                            </SouthBoundingCoordinate>
                        </BoundingRectangle>
                    </Geometry>
                </HorizontalSpatialDomain>
            </Spatial>
        </Granule>
    </Granules>
</GranuleMetaDataTable>
```

```
</HorizontalSpatialDomain>
</Spatial>

<Platforms>
    <Platform>
        <ShortName>Aura</ShortName>
        <Instruments>
            <Instrument>
                <ShortName>HIRDLS</ShortName>
                <Sensors>
                    <Sensor>
                        <ShortName>HRDLS Channel 21</ShortName>
                    </Sensor>
                </Sensors>
            </Instrument>
        </Instruments>
    </Platform>
</Platforms>

<Campaigns>
    <Campaign>
        <ShortName>HRDLS</ShortName>
    </Campaign>
</Campaigns>

<Price>10.00</Price>

<OnlineAccessURLs>
    <OnlineAccessURL>
        <URL>
            ftp://airsparlu.ecs.nasa.gov//AIRS.G07270040936.hdf
        </URL>
        <URLDescription>Description</URLDescription>
    </OnlineAccessURL>
</OnlineAccessURLs>

<OnlineResources>
    <OnlineResource>
        <URL>The Resource</URL>
        <Description>The Description</Description>
        <Type>Text</Type>
    </OnlineResource>
</OnlineResources>

<Orderable>false</Orderable>

<DataFormat>HDF</DataFormat>

<CloudCover>40.1</CloudCover>

<AssociatedBrowseImages>
    <ProviderBrowseId>AIRS.2005.G07239125629.jpg</ProviderBrowseId>
</AssociatedBrowseImages>
</Granule>
</Granules>
</GranuleMetaDataTable>
```

4. BROWSE METADATA

4.1. Required Fields

Table 5. Browse Required Fields

Name	Data Type	Constraint	Description
ProviderBrowseId			The unique id for the browse image given by the provider. This should be unique per provider.
BrowseImageFileName			The name of the delivered browse image
BrowseImageFileSize			The size of the delivered browse image in bytes. A browse image delivered that is a different size than indicated will be rejected.

4.2. Recommended Fields

Table 6. Browse Recommended Fields

Name	Data Type	Constraint	Description
InsertTime			This attribute specifies the date/time this browse image was created in the data provider's database.
LastUpdate			This attribute specifies the last time the browse image was modified in the provider's database.

4.3. Sample Metadata File

```
<?xml version="1.0" encoding="utf-8"?>
<BrowseMetaDataTable>

    <BrowseImages>
        <BrowseImage>
            <ProviderBrowseId>AIRS.2005.G07239125629.hdf.jpg</ProviderBrowseId>
            <InsertTime>2007-10-21T05:00:00Z</InsertTime>
            <LastUpdate>2007-10-31T14:53:00Z</LastUpdate>
            <BrowseImageFileName>
                AIRS.2005.G07239125629.hdf.jpg
            </BrowseImageFileName>
            <BrowseImageFileSize>118660</BrowseImageFileSize>
        </BrowseImage>
    </BrowseImages>
</BrowseMetaDataTable>
```